

An unlevel playing field: national income estimates and reciprocal comparison in global economic history

Morten Jerven

School for International Studies, Simon Fraser University, 7200-515 West Hastings Street, Vancouver, BC, V6B 5K3, Canada

E-mail: mjerven@sfu.ca

Abstract

If we take recent income per capita estimates at face value, they imply that the average medieval European was at least five times ‘better off’ than the average Congolese today. This raises important questions regarding the meaning and applicability of national income estimates throughout time and space, and their use in the analysis of global economic history over the long term. This article asks whether national income estimates have a historical and geographical specificity that renders the ‘data’ increasingly unsuitable and misleading when assessed outside a specific time and place. Taking the concept of ‘reciprocal comparison’ as a starting point, it further questions whether national income estimates make sense in pre-and post-industrial societies, in decentralized societies, and in polities outside the temperate zone. One of the major challenges in global history is Eurocentrism. Resisting the temptation to compare the world according to the most conventional development measure might be a recommended step in overcoming this bias.

Keywords Africa, economic history, national income accounting, reciprocal comparison

A recent surge in interest in the long-term global distribution of economic growth has increased the demand for historical national income estimates. Comparable measures of welfare, productivity, and income through history are potentially helpful for scholars wishing to confirm old or form new hypotheses on long-term patterns of divergence and convergence. It is desirable to meet the demand for a quantitative resolution of these big questions – including, for example, the ‘great divergence’ as invoked by Kenneth Pomeranz – with innovative estimates.¹ This is particularly true for cliometricians, whose investigation is effectively limited by the availability of quantitative evidence. This demand has been met

1 Kenneth Pomeranz, *The great divergence: China, Europe, and the making of the modern world*, Princeton, NJ: Princeton University Press, 2000. For one recent attempt at quantitative resolution see Tirthankar Roy, ‘Economic conditions in early modern Bengal: a contribution to the divergence debate’, *Journal of Economic History*, 70, 1, 2010, pp. 179–94; see also the papers deriving from the project ‘Historical patterns of development and underdevelopment: origins and persistence of the great divergence’, <http://www.cepr.org/research/HIPOD.htm> (consulted 19 September 2011).

by the work of a range of scholars, most prominently the late Angus Maddison, who provided new national account estimates for countries and periods that were previously unmeasured.²

In theory, GDP, or national income accounting, is a measure of the total value of all economic activities within a nation, representing the combined ability of the economic actors within a given area to produce and consume goods and services.³ In practice, adequate approximation is constrained by data collection methods and by the availability of the required data. In addition, it is far easier to define production and consumption in a general manner than to apply this definition universally to the actual economic transactions that have taken place throughout history so as to standardize the measure across time and space. Scholars have generally acknowledged these theoretical and practical difficulties, but the spatial and temporal constraints of the measure, resulting in systematic measurement errors, are less recognized and studied.

This article disputes the claim to universality of the national income concept in practice, and suggests that to correctly assess its applicability the measures need to be fully historicized and contextualized.⁴ The GDP measure should not be taken at face value, because it does not fit with its theoretical definition. In practice the measure can be used as an approximation, but its usefulness is constrained by empirical applicability. This investigation focuses on the importance of these constraints throughout time and space, and its main question is whether this variation results in a serious bias in the income level estimates. It should be remembered that GDP estimates play a dual role in global economic history. In some contexts they are ‘data’, as per the original meaning – ‘something that is given’ – while in other contexts they are scholarly ‘products’.⁵ Particularly since the 1960s, GDP estimates have been provided and produced by national statistical services as official estimates of national income. These estimates are quite different from the ones that are independently produced or extrapolated to apply to periods of time and geographical locations that lack official statistical services geared towards the production of this metric. In the latter situation, the GDP estimates are ‘products’, the result of a ‘model’ for which the assumptions and fit to empirical context and the data basis can and should be interrogated.

The standard objection from an economist to issues of the unreliability of these measures would be that, if a number or a statistical result is based on enough observations, any errors in the constituent data will balance each other out. This standard defence of inferential statistics does not apply here: these errors are not just random; they are systematic. Thus comparing places and times as different as medieval Europe and contemporary Africa in

2 Angus Maddison, *Contours of the world economy, 1–2030 AD: essays in macro-economic history*, Oxford: Oxford University Press, 2007. Maddison’s work continues to be revised and expanded under the auspices of the ‘Maddison Project’: see <http://www.ggdc.net/maddison/maddison-project/> (consulted 9 December 2011).

3 For a basic introduction to the technicalities of the measure, see François Lequiller and Derek Blades, *Understanding national accounts*, Paris: OECD, 2007.

4 A premise shared with a article recently published in this journal: Daniel Speich, ‘The use of global abstractions: national income accounting in the period of imperial decline’, *Journal of Global History*, 6, 1, 2011, pp. 7–28.

5 G. Austin, ‘The “reversal of fortune” thesis and the compression of history: perspectives from African and comparative economic history’, *Journal of International Development*, 20, 8, 2008, pp. 996–1027.

terms of GDP per capita is misleading. The two places are quantitatively incommensurable on a scale large enough to affect the debates on patterns of divergence and convergence in global economic history.⁶

This article will first historicize the national income concept, and then suggest possible demarcation points in time for its usage. It argues that it may be misleading to extrapolate ‘GDP’ as a development criterion to times when the ‘GDP’ was either consciously or subconsciously rejected, and perhaps even more so to polities that did not yet exist. The article examines differences created by space as well as time, arguing that certain aspects of geography or location may be considered ‘endogenous’ to the GDP measure, which means that they have a direct impact on the measurement itself and may bias analysis. Finally, it argues that the GDP measure is, in essence, a tool of a centralized state planning for development. Centralized states have a stronger ability and motivation to record data on production than decentralized states or loosely organized polities, but the ability of the state to record and tax production is not necessarily positively correlated with higher standards of living.

These three lines of argument are brought together in the conclusion. If we are to take the principle of ‘reciprocal comparison’ seriously in global economic history, and also seek to avoid ‘Eurocentrism’,⁷ an exclusive reliance on national income level estimates must be reconsidered. Pomeranz has argued that, rather than abandoning comparisons altogether, one should aim at producing better ones. The key is to make reciprocal comparisons: one should look at deviations on both sides of the comparison, and view both sides as the norm, rather than letting one method or institution be the standard by which other societies and places are measured. Accepting this line of reasoning, this article agrees that GDP estimates may be helpful for certain purposes, but when applied to global economic history they need to be fully historicized and contextualized. The article provides some insight into how this may be done.

Global history, reciprocal comparisons, and historical national accounts

In a very specific sense, Angus Maddison was correct to assert that ‘quantification clarifies issues which qualitative analysis leaves fuzzy. It is more readily contestable and likely to be contested. It sharpens scholarly discussion, sparks off rival hypotheses, and contributes to the dynamics of the research process.’⁸ This may be an appropriate summary of the mainstream consensus. It is also true that quantitative evidence has a comparative advantage over qualitative evidence, but this is because of the former’s *perceived* ease of

6 This article will focus on GDP estimates. Others have explored the use of alternative measures, most notably real wages, which may also have significant biases. For a review of recent contributions in this direction, see Anne Booth, ‘Review article: living standards in the past: new perspectives on well-being in Asia and Europe’, *Journal of Global History*, 1, 2, 2006, pp. 289–92.

7 The principle gained currency with the work of R. Bin Wong, *China transformed: historical change and the limits of European experience*, Ithaca, NY: Cornell University Press, 1997, was forcefully restated by Pomeranz, *Great divergence* (in particular pp. 8–10), and was more recently adapted to the study of Africa by Gareth Austin, ‘Reciprocal comparison and African history: tackling conceptual Eurocentrism in the study of Africa’s economic past’, *African Studies Review*, 50, 3, December 2007, pp. 1–28.

8 Maddison, *Contours*, p. 1.

interpretation, not its analytical value. As a result, a statement such as ‘the informal sector was of great importance in the Kenyan economy in the 1980s’ is likely to have less scholarly traction than a paper claiming that ‘the informal sector accounted for 77% of the Kenyan economy in 1986’, even though the latter statement is invalid by definition. The most commonly accepted definition of the informal economy is that it is unrecorded; thus we do not have any direct data on the sector that could assert its size without a certain level of inaccuracy. So the question ‘how important was the informal sector in Kenya in 1980s’ can be ‘clarified’, but it comes at a cost; and whether it is intellectually defensible to provide potentially misleading and ‘fuzzy’ numbers is a matter of academic judgement.

National income estimates are not facts but products, and most obviously so outside the period of official national accounting. In the exchange of opinions published in the 2008 *European Review of Economic History*, Gregory Clark’s interpretation of the pre-1800 world as ‘Malthusian’ was challenged by a range of scholars.⁹ Karl Gunnar Persson, in particular, made the point that Clark’s analysis was not compatible with Maddison’s estimates.¹⁰ Clarke rebuffed the criticism by writing that he ‘dismissed Maddison’s estimates not because they are inconvenient to Malthusian theory, but because they are based on nothing more than Maddison’s incorrect assumptions about how economies worked before 1800. They are not based on any serious historical evidence’.¹¹ There is now a considerable scholarly effort underway to improve and document the historical evidence that underlies the national income estimates for Europe and other regions before 1900, and this research tends to revise the estimates upwards.

One should be careful when using national income estimates to compare achieved levels of development, progress, organizational sophistication, and welfare across time and space. Alexander Gerschenkron offers a precaution, suggesting that, if estimates from quantitative data ‘are at variance with what we should expect from general historical knowledge, they should be rejected’.¹² There is a major problem here. There are many questions to which ‘general historical knowledge’ does not give a straightforward answer. A case in point is the ‘great divergence’. In recent years the uniqueness of Europe and the first industrial revolution has been called into question,¹³ especially by historians of China, who argue that the parts of Europe that industrialized did not have a quantitative advantage vis-à-vis parts of China around 1750.¹⁴ In addition, some historians of Europe, most prominently Nicholas Crafts, challenge the mainstream view that the English industrial revolution led to a pronounced spurt of growth. Crafts has revised growth accounts to establish that the

9 *European Review of Economic History*, Symposium on Gregory Clark’s *A Farewell to Alms*, 12, 2, 2008.

10 Karl Gunnar Persson, ‘The Malthus delusion’, *European Review of Economic History*, 12, 2, pp. 165–73, esp. p. 166.

11 Gregory Clark, ‘In defense of the Malthusian interpretation of history’, *European Review of Economic History*, 12, 2, pp. 175–99, esp. p. 177.

12 A. Gerschenkron, *Economic backwardness in historical perspective: a book of essays*, Cambridge, MA: Belknap Press of Harvard University Press, 1962, quoted in D. C. M. Platt, *Mickey Mouse numbers in world history: the short view*, Basingstoke: Macmillan, 1989, p. 1.

13 Patrick K. O’Brien, ‘The reconstruction, rehabilitation and reconfiguration of the British Industrial Revolution as a conjuncture in global history’, *Itinerario*, 24, 2000, pp. 117–34.

14 Pomeranz, *Great divergence*, pp. 29–68, makes the case for the levels of material advancement in Europe and Asia as having ‘surprising resemblances’.

industrial revolution was *not* associated with rapid aggregate growth, as was commonly assumed.¹⁵ Other scholars vehemently argue for the superiority of Europe, and the accumulated weight of ‘general historical knowledge’ would probably lean towards European superiority.¹⁶ But some would argue that this weight is the result of a long tradition of Eurocentric scholarship, going back at least as far as Max Weber and Karl Marx, rather than accurate data.

The most desirable feature of numbers is that they may stimulate new and vigorous academic debate. However, the historical national accounts have tended to reinforce rather than challenge dominant patterns. A popular view of the non-European world is that it was stagnant and void of economic change and dynamism. Currently, world income tables support such an interpretation of the African continent’s history.¹⁷ According to Maddison’s estimates, African GDP per capita for the years 1500, 1600, 1700, and 1820 was 414, 422, 421, and 420 respectively.¹⁸ Thus civilizations rose and fell in parts of Africa, millions of slaves were sold and purchased in the transatlantic slave trade, and the ‘cash crop revolution’ took place, but the national income statistics barely blinked in response.¹⁹ A lack of data is partly to blame here, but so is the lack of scholarly work focused on creating comparable accounts of the data that does exist in different sources.²⁰

By contrast with Africa, there are far more hands at work in European economic history, and some of these have been active in evaluating data from eras that lack the statistics produced by modern governments. This research has led to revisions in various measures of income. Some of the recent work on medieval income estimates for Europe, for example, has revised income per capita significantly upwards as seen in Table 1.²¹ The ‘old’ Maddison estimates for the year 1500 were significantly lower: 75% in the case of England, and 90%, 50%, and 95% lower in the case of Holland, Italy, and Spain respectively.

If we compare the data on pre-modern economies with the data on some contemporary poor countries as shown in Table 2, we find that England in the year 1270 was three times

15 N. Crafts, ‘Productivity growth in the Industrial Revolution: a new growth accounting perspective’, *Journal of Economic History*, 64, 2004, pp. 521–35.

16 For a classic statement of European superiority, see Eric L. Jones, *The European miracle: environments, economies and geopolitics in the history of Europe and Asia*, Cambridge: Cambridge University Press, 1981; D. Landes, *The wealth and poverty of nations: why some are so rich and some so poor*, New York, NY: W. W. Norton & Company Inc., 1998; D. North, *Institutions, institutional change and economic performance*, Cambridge: Cambridge University Press, 1990.

17 Maddison, *Contours*.

18 These data include North Africa and are in 1990 international dollars. The data for ‘Sahel and West Africa’ and ‘Rest of Africa’ are completely stagnant at 415 dollars throughout those years. *Ibid.*, p. 185.

19 There is some evidence of cash-crop-related growth for Ghana. Historical estimates are available for 1891, 1901, and 1911, and these show rapid growth in the period. For further discussion, see Morten Jerven, ‘African growth recurring: an economic history perspective on African growth episodes, 1690–2010’, *Economic History of Developing Regions*, 25, 2, 2010, pp. 127–54.

20 Patrick Manning, ‘The prospects for African economic history: is today included in the long run?’, *African Studies Review*, 30, 1987, pp. 49–62.

21 See recent new estimates for medieval Britain in Stephen Broadberry, Bruce Campbell, Alexander Klein, Mark Overton, and Bas van Leeuwen, ‘British economic growth, 1270–1870’, unpublished paper presented at the Third European Congress on World and Global History, London School of Economics, April 2011, <http://www2.lse.ac.uk/economicHistory/pdf/Broadberry/BritishGDPLongRun.pdf> (consulted 9 December 2011), p. 61.

Table 1. GDP per capita levels (in 1990 international dollars) for medieval and early modern Europe.

Year	England	Holland	Italy	Spain
1270	638			
1300	739		1644	
1348	803	876		
1400	1303	1195	1726	
1500	1128	1454	1644	1295
1600	1126	2662	1302	1382
1700	1568	2105	1398	1230

Source: Broadberry et al., 'British economic growth', p. 61.

Table 2. GDP per capita levels (in 1990 international dollars) for China, India, Congo, and Ghana in the twentieth century.

Year	China	India	Congo	Ghana
1913	552	673		781
1950	448	619	570	1122
2000	3421	1892	218	1265

Source: Maddison Project.

'better off' or 'more developed' than Congo in the year 2000. In the year 1950, India was as developed as England was in the year 1270, and Ghana, considered a 'middle-income' poor country today, has still to reach medieval European standards of living.²² These comparisons bring into question the suitability of using this particular measure of development across time and space. If one described the standard of life in an African middle-income country today as medieval, one would in most forums be labelled as prejudiced, or insensitive to local conditions. Yet, this is what the data are telling us. Does the use of national income estimates in global comparative economic history therefore make any sense? Although many countries today have national statistical offices, this development might not in itself provide enough of a basis for an objective comparison. GDP as a concept had specific historical and geographic origins, which might well limit its universal application to global economic history.

Historicizing national income: from Richard Stone to Mahbub ul Haq

The concept of national income is as old as economic thinking itself. Its origins go back to William Petty (1623–87) and to François Quesnay (1694–1774) and his *Tableau*

22 The lowest income recorded in the Maddison dataset is for Sri Lanka in 1820, at 83 dollars per capita. The second lowest was recorded for Congo (Kinshasa) in 2001, at 207 dollars.

Economique.²³ Its institutionalized history, however, is much shorter and set within a specific historical period and geographical region. National income accounting was developed for the industrial, centralized, modern nation-states in the northern hemisphere during the interwar period. In 1928, the International Convention Relating to Economic Statistics provided the legal basis for founding an international statistical office under the auspices of the League of Nations.²⁴ In 1945, Richard Stone chaired the Sub-committee on National Income of the League of Nations and drafted a memorandum, published two years later under the title 'Measurement of national income and the construction of social accounts'. In 1952, Stone developed the System of National Accounts (SNA) at the Organization for European Economic Cooperation, published one year later as the United Nations System of National Accounts.²⁵ Since then, the standard of national accounts has been updated and revised in 1968, 1993, and 2008 but, 'although they pay lip service to the subsequent revisions ... many countries still adhere to the basic system and its corresponding accounting foundations as first set out'.²⁶

The main features of this system reflect the priorities and characteristics of the time and place in which the measure of national income was devised. The directors and the statisticians on the committee were almost all from Western and industrialized countries (Japan being the notable exception to the first criterion), and the classifications reflected the system of trade and production prevalent in these nations. In particular, they were suitable for developed economies in which most economic activity was formalized and recorded, while little attention was given to issues pertaining to less developed economies, where economic activities were characterized by informality and lack of recording.

Some analysts recognized and critiqued the limitations of this system almost as soon as it had been developed. Dudley Seers, for example, reminded data users that 'what appear to be merely technical choices in statistics are in fact often of profound importance, because published data mould our perception of reality'.²⁷ He questioned several basic assumptions of the SNA. The first of these was the assumption that the economy could be considered monistic; many scholars, including W. Arthur Lewis, considered dual models as more appropriate for less developed economies.²⁸ A second assumption was that nations were economically independent, which was not the case in countries with significant foreign-owned industries. A third assumption was that capital formation was fixed, and could be

23 Randall G. Holcombe, 'National income accounting and public policy', *Review of Austrian Economics*, 17, 4, 2004, pp. 387–405.

24 Michael Ward, *Quantifying the world: UN ideas and statistics*, Bloomington, IN: Indiana University Press, 2004.

25 Organisation for European Economic Cooperation, *A standardized system of national accounts*, OEEC: Paris, 1952; United Nations Statistical Office, *A system of national accounts and supporting tables*, UN: New York, 1953.

26 Ward, *Quantifying the world*, p. 45.

27 Dudley Seers, 'The political economy of national accounting', in Alec Cairncross and Mohiner Puri, eds., *Employment, income distribution and development strategy: problems of the developing countries*, New York: Holmes and Meier, 1976, p. 193. Seers first published his criticism of the system in Dudley Seers, 'The role of national income estimates in the statistical policy of an under-developed area', *Review of Economic Studies*, 20, 3, 1952–53, pp. 159–68.

28 See, among others, W. A. Lewis, 'Economic development with unlimited supplies of labour', *Manchester School* 1954, 22, 2, 1954, pp. 139–91.

defined and measured using a limited number of metrics. Seers pointed out that investments in urban real estate are captured by this measure, but investments in agricultural land and rural infrastructure are not. One could add to this that investments in what we today call ‘human capital’ were also missing. A fourth assumption was that market and non-market production could be clearly differentiated. Related to that assumption is the most important assumption of them all: that reliable statistics were available.²⁹ Seers thus argued that the SNA had limited applicability to developing countries, that its assumptions shaped how we thought about and perceived development, and that the system was first and foremost designed with developed countries in mind.

While the SNA was not itself at fault, problems arose in terms of how it was put to use in policy planning and evaluation. The concept of GDP and its operationalization in the SNA strengthened the Keynesian paradigm in the 1950s and 1960s, which was further strengthened by the influential Harrod–Domar growth model. The basic premise of the Harrod–Domar growth model is that the rate of growth in national income is determined by the capital–output ratio. This component was in turn determined by the savings ratio; thus the key factor for growth was the mobilization of domestic and foreign savings, as emphasized by Walt Rostow in his *Stages of economic growth*.³⁰ In this conceptualization of development, growth is a function of fixed capital formation, which was expected to come from the industrial sector. These theories formed the basis of development policy in India during the period of the second five-year plan from 1956 to 1961, when, under the oversight of Prasanta Chandra Mahalanobis, heavy industry was prioritized.³¹ Furthermore, this intuition became central to the ruling development paradigm until the 1980s and, as argued by William Easterly, served as the basic justification for official development assistance as overseen by international financial institutions such as the World Bank.³² Thus the SNA and the related input–output tables clearly shaped the development process, in which certain economic activities were valued more highly than others.

In particular, one striking absence from both national income estimations and the Harrod–Domar growth model is the role of services. Wassily Leontieff, one of the pioneers in implementing national accounting and input–output modelling for economic analysis, commented that, ‘among the lines of economic activity completely ignored in our analysis, the most important are the entire fields of (a) distribution, wholesale and retail, (b) banking and finance, and (c) all non-rail transportation’.³³ In addition, Leontieff noted that his calculations omitted ‘the income–expenditure accounts of all public bodies, including the budgets of federal, state and local governments’.³⁴ For Leontieff, ignoring services and public expenditures was acceptable, and he was satisfied that ‘the relation of all the

29 Seers, ‘Political economy’, pp. 195–9.

30 W. W. Rostow, *Stages of economic growth: a non-communist manifesto*, Cambridge: Cambridge University Press, 1960.

31 V. N. Balasubramanyam, *The economy of India*, London: Weidenfeld and Nicolson, 1984.

32 W. Easterly, ‘The ghost of financing gap: testing the growth model used in the international financial institutions’, *Journal of Development Economics*, 60, 1999, pp. 423–38.

33 Wassily W. Leontieff, *The structure of the American economy, 1919–1929: an empirical application of equilibrium analysis*, Cambridge, MA: Harvard University Press, 1941, pp. 21.

34 *Ibid.*

unaccounted economic units to the rest of the system is implicitly reflected in the anonymous undistributed account'.³⁵ In this model, services were not considered to play an important role in total productivity.

Leontieff's focus on physical production alone may have worked for the mid-twentieth-century European or American economy, but this approach does not work well today. Because services are now a more important part of the economy, they have come to take a more central place in productivity investigations.³⁶ The case of recent economic growth in India is illuminating. According to the International Standard Industrial Classification (ISIC), the impressive recent growth in India is classified as originating primarily in the service sector, but an emerging question is whether we should understand and conceptualize the 'service' growth in India as actually quite similar to 'manufacturing' growth in other developing countries or in other time periods.³⁷ The information- and communication-technology-based service industry in India has many characteristics in common with what we expect to see in modern industry: the central role of technology, economies of scale, and a global market. It would still not be included as manufacturing in standard SNA measures, however. When SNA travels outside the time and place in which it was developed, it can create a mismatch of categories and the neglect of niches of dynamism, which can distort economic analysis of earlier historical periods as much as that of the contemporary world.³⁸ Thus the example of contemporary India suggests that this system is not a very useful guide for detecting changes in productivity and living standards across time and space.

In 1990, the United Nations published the first *Human Development Report*, which launched a new primary indicator of development, the Human Development Index. This step was explicitly motivated by the increasingly widely perceived need to counter the dominance of GDP as an indicator of development.³⁹ The pioneering leader of this approach was Mahbub ul Haq, who, together with Amartya Sen, argued for the institutionalization of this indicator of development instead of other candidates such as 'basic needs' or 'quality of life'.⁴⁰ In the previous paradigm, in which GDP predominated, 'economic growth' was the variable targeted by policy. Importantly, in the system of national accounts, 'capital formation' did not include investment in human capital. Rather it was classified as unproductive government expenditures. In contrast, the new 'capabilities approach' given in the rationale for the Human Development Index explicitly acknowledged that the aims of

35 *Ibid.*

36 Stephen Broadberry, *The productivity race: British manufacturing in international perspective, 1850–1990*, Cambridge: Cambridge University Press, 1997.

37 For a discussion of the categorization of recent Indian growth, see for instance Rashmi Banga, 'Critical issues in India's service-led growth', Indian Council for Research on International Economic Relations (ICRIER) Working Paper 171, October 2005, <http://www.icrier.org/pdf/WP171.pdf> (consulted 9 December 2011).

38 David Washbrook notes the strong standing of the service sector in pre-modern India in 'Colonialism, globalization and the economy of South-east India, c.1700–1900', Working Papers of the Global Economic History Network (GEHN) 24/06, 2006, eprints.lse.ac.uk/22466/1/wp24.pdf (consulted 9 December 2011).

39 UNDP, *Human Development Report*, Oxford: Oxford University Press, published annually since 1990.

40 Amartya Sen, 'Human Development Index', in David A. Clark, ed., *The Elgar companion to development studies*, Cheltenham: Edward Elgar, 2006, pp. 256–60.

development should be gains in education and health, rather than simply hoping that these would result from economic growth.

This turn away from economic growth as the primary signifier of development has continued in two waves. First, IMF- and World Bank-initiated structural adjustment programmes went from primarily targeting economic growth to emphasizing poverty reduction.⁴¹ Second, the UN adopted the Millennium Development Targets as the coordinating development principles for both national governments and international donors, thus further reducing the importance of economic growth and national income accounting as development indicators.⁴² In the academic domain, using different indicators of development leads scholars to reach conclusions that are in conflict with one another, as neatly summed up by the contesting claims of Pritchett ('Divergence, big time') versus Kenny ('Why are we worried about income? Nearly everything that matters is converging').⁴³

GDP measures a very specific concept of development: the growth of large-scale industrial production and the reach of the nation-state, defined as its ability to tax and record marketed production. This is not always what we are interested in measuring when we speak of development, however, and it adds a teleological perspective to economic development, which has made it open to criticism.⁴⁴ GDP is thus valid for making global comparisons only in the limited period that it has explicitly been the main instrument for states to monitor their own progress and development. The League of Nations' international convention in 1928 and the UN's publication of the Human Development Report in 1990 might be seen as appropriate beginning and end points. Before and after these dates the interpretation of national income as 'development' becomes less straightforward, making it less useful as a concept for understanding progress and development in both pre- and post-industrial societies.

The widespread use of GDP as an indicator of development also has implications for how we use the measure as evidence in economic history. We are often interested in whether a specific economic policy was successful, or, more generally, whether a particular strategy was efficient, considering the aims that were set out and the means that were available. In global history, however, GDP can be used to measure this only some of the time. Effective reciprocal comparison thus recognizes that development may take different forms, and is even more likely to do so from different starting points in time.⁴⁵

41 J. E. Stiglitz, 'Towards a new paradigm for development: strategies, policies, and processes', Prebisch Lecture at UNCTAD, Geneva, 1998, <https://ceaemgmt.colorado.edu/ceae/images/File/mcedc/prebisch98.pdf> (consulted 9 December 2011).

42 Richard Jolly, 'The MDGs in historical perspective', *IDS Bulletin*, 41, 1, January 2010, pp. 48–50.

43 Lant Pritchett, 'Divergence, big time', *Journal of Economic Perspectives*, 11, 3, 1997 pp. 3–17; Charles Kenny, 'Why are we worried about income? Nearly everything that matters is converging', *World Development*, 33, 2005, pp. 1–19.

44 Rostow's *Stages of economic growth*, which hails from the same period and place as the national income estimation method, does this as well. Easterly, in the 'The ghost of financing gap', is one of many critics of this view of development.

45 The most famous statement of the variation across the time axis was simply that 'late development will be different', paraphrasing Gerschenkron, *Economic backwardness*.

Nature, climate, and factor endowments: implications for measuring development

Temporal differences in development paths and paradigms have become widely recognized, but spatial differences less so. This section explores the problems of comparing the value of production and standards of living across radically different regions of the world. Although many scholars and policy makers recognized that national income had outplayed its role as the main instrument and aim of economic policy by the 1990s, it is still the most important piece of evidence employed by economists and economic historians today. One important reason for its continued scholarly dominance is the launch of the Penn World Tables in the late 1980s, which have been updated on a regular basis ever since their creation.⁴⁶ This major statistical event made cross-country comparisons possible, and it attempted to solve the most basic problems that arise when comparing consumption and production across space.

The national income data are aggregated in local currencies. The first step to take towards making the figures comparable is to use foreign exchange rates in order to express one country's income in the currency of another. This does not take care of the problem of differences in domestic prices on non-tradable goods, however, or other factors that cause a divergence in purchasing power parity (PPP).⁴⁷ To achieve PPP one needs to adjust for the fact that one dollar goes a lot further in Ethiopia than it does in Canada. This entails a complicated process of collecting prices and then determining a basket of goods and services to weight the individual prices.⁴⁸ Price data and baskets are regularly updated under the auspices of the International Comparison Program.⁴⁹ Lack of comparable information may introduce bias in these comparisons; it may well be the case that in Ethiopia, for example, prices are more easily available for urban areas than for rural areas, and furthermore that there may be better price data on imported goods than on domestically produced subsistence goods. If the urban data on imported goods are weighted too heavily in the consumption baskets, the PPP-adjusted GDP figures will understate the living standard in Ethiopia.

Using the national income data to project backwards creates still more problems, as very often the projected income levels are too low. Debin Ma and co-authors have re-estimated benchmark PPP years for 1936 in Japan, China, Korea, and Taiwan, and found that previous level estimates by Madison, based on a backwards projection from 1990, present income levels of these economies that are too low.⁵⁰ The biggest controversy at the moment

46 Nicholas Stern, 'The economics of development: a survey', *Economic Journal*, 99, 597, 1989, p. 685.

47 Alan M. Taylor and Mark P. Taylor, 'The Purchasing Power Parity debate', *Journal of Economic Perspectives*, 18, 4, 2004, pp. 135–58.

48 Angus Deaton and Alan Heston, 'Understanding PPPs and PPP-based national accounts', *American Economic Journal: Macroeconomics*, 2, 4, 2010, pp. 1–35.

49 The last round was in 2005, and the next round is planned to finalize in 2011. See http://siteresources.worldbank.org/ICPEXT/Resources/ICP_2011.html (consulted 9 December 2011).

50 Debin Ma, Kyoji Fukao, and Tangjun Yuan, 'Real GDP in pre-war East Asia: a 1934–36 benchmark Purchasing Power Parity comparison with the U.S.', *Review of Income and Wealth*, 53, 3, 2007, pp. 503–37.

is the accuracy of the level estimates of Chinese GDP.⁵¹ Surjit S. Bhall has argued that data on the income level and the growth rate cannot be right, as it would place the average income level in China in the 1950s at a level too low to sustain its population.⁵²

Despite these problems, the availability of these PPP-adjusted figures allowed for a surge in the empirical growth literature, which Durlauf et al. refer to as the ‘growth regression industry’.⁵³ The industrious comparisons of income levels and growth rates across time and space found in this literature might still be hazardous, however, and in particular the data for poor countries is unreliable.⁵⁴ Exactly how misleading these comparisons are is still an open question, precisely because we do not have accurate information on what a ‘correct’ world income distribution would look like.⁵⁵ Some indications of the problems involved have been derived from comparing the different versions of the Penn World Tables.⁵⁶ It has been concluded that the variation in growth rates between the different datasets makes the analysis of short-term growth in non-OECD countries particularly risky,⁵⁷ and that different datasets support different determinants of growth when used in econometric modelling.⁵⁸ The differences in price comparisons are particularly transparent when compiling absolute poverty measures, such as estimating the share of the world population living on less than a dollar a day. Deaton reports that the 2005 revision meant an increase of half a billion people under the poverty line.⁵⁹ These discrepancies and inaccuracies for income data between 1960 and today are serious because they are likely to be systematic rather than random. Furthermore, there are good reasons to think that the problems are even greater once we compile data to evaluate global development trajectories across centuries.

In a rational choice interpretation of history, the choice of technology, defined as innovation rather than invention, was typically conditioned by the environment. An

51 Robert C. Feenstra, Hong Ma, C. Peter Neary, and D. S. Prasada Rao, ‘How big is China? and other puzzles in the measurement of real GDP’, University of California at Davis, March 2010, http://www.econ.ucdavis.edu/faculty/fzfeens/pdf/Feenstra_Ma_Neary_Rao_2010_Oxford.pdf (consulted 9 December 2011).

52 Surjit S. Bhalla, ‘World Bank: most Asians dead in 1950’, *Business Standard*, 23 August 2008.

53 Steven N. Durlauf, Paul A. Johnson, and Jonathan R. W. Temple, ‘Growth econometrics’, in Philippe Aghion and Steven N. Durlauf, eds., *Handbook of Economic Growth*, Amsterdam: Elsevier, 2005, pp. 555–667.

54 John W. Dawson, Joseph P. DeJuan, John J. Seater, and E. Frank Stephenson, ‘Economic information versus quality variation in cross-country data’, *Canadian Journal of Economics*, 34, 3, 2001, pp. 988–1009.

55 For a study of different distributions of African GDP, see Morten Jerven, ‘The relativity of poverty and income: how reliable are African economic statistics?’ *African Affairs*, 109, 434, 2010, pp. 77–96.

56 A comparison of the growth rates in Penn World Tables, Maddison, and World Development Indicators for a sample of African economies was made by Morten Jerven, ‘Random growth in Africa? Lessons from an evaluation of the growth evidence on Botswana, Kenya, Tanzania and Zambia, 1965–1995’, *Journal of Development Studies*, 46, 2, 2010, pp. 274–94.

57 Simon Johnson, William Larson, Chris Papageorgiou, and Arvind Subramanian, ‘Is newer better? Penn World Table revisions and their impact on growth estimates’, National Bureau of Economic Research Working Paper 15455, 2009, http://baselinescenario.files.wordpress.com/2009/10/jlps_oct-2009-final.pdf (consulted 9 December 2011).

58 Antonio Ciccone and Marek Jarocinski, ‘Determinants of economic growth: will data tell?’, *American Economic Journal: Macroeconomics*, 2, 4, 2010, pp. 223–47.

59 Deaton, Angus, ‘Price indexes, inequality, and the measurement of world poverty’, *American Economic Review*, 100, 1, 2010, pp. 5–34.

illustrative example is the wheel, for only in places where there were draught animals and a landscape that allowed the building of wide roads did it emerge as an important new technology. This is one of the lessons to be taken from research on pre-industrial societies. The implication is that the rationality of choices regarding technology or production techniques is dependent on the conditions under which these choices were made. In his economic history of pre-colonial West Africa, Antony Hopkins noted that ‘Comparing the natural resources and climates of different parts of the world in order to draw conclusions about whether they stimulated or retarded the economic progress of particular societies is a tempting but unprofitable exercise – rather like trying to decide if life is more difficult for penguins in the Antarctic or camels in the Sahara.’⁶⁰ Issues such as the choice of production techniques and the investment in physical and human capital need to be evaluated within specific environments and local conditions.

To start thinking about this issue, an example based on broadly accepted historical notions – so-called ‘stylized facts’ – is helpful. Let us compare the ‘north’, exemplified by medieval England, with the ‘south’, exemplified by pre-colonial Asante. In medieval England there was limited land available. Climate (particularly having to survive the winter) and soil fertility determined when planting would occur and what types of crops would be planted in order to satisfy nutritional requirements. The grain harvest in England was typically an annual harvest. This harvest was in turn subject to tithe, and parts of the harvest went to breweries and mills. Meanwhile, in Asante, land was abundant but holding livestock was impossible because of the prevalence of the tsetse fly that spread sleeping sickness (trypanosomiasis).⁶¹ This in turn limited the availability of manure for fertilizer and the use of the plough. In the tropical forest belt, peasants can rely on food crops such as banana and tubers. These crops are harvested on a day-to-day basis, which means that the food harvest was, and still is, immeasurable for all practical purposes.

These comparative differences have implications for the measurement of national income. The climatic conditions in the ‘north’ allowed for storage of food and therefore provided relatively strong incentives and opportunities for refining and processing food-stuffs. Furthermore, holding livestock requires more input to produce the same amount of calories than raising cereals or other food crops. In national income terms, adaptation to the geographical conditions in England led to more expensive or ‘valuable’ production and consumption than that of pre-colonial Asante. The nature of the harvesting cycle made it easier to record an annual estimate of total food production and the extent of forward linkages for English agricultural production, thus further increasing the national income estimates of pre-modern England upwards.

It may be argued that some of these geographical conditions are root causes of current differences in development.⁶² But it should be remembered that this is only one of the

60 Antony G. Hopkins, *An economic history of West Africa*, Longmans: London, 1973, pp. 13–14.

61 The lack of cattle due to tsetse-borne trypanosomes, causing trypanosomiasis, is specific to the African tropics. Non-tsetse-borne trypanosomes existed elsewhere in the tropics, but essentially affected equids rather than bovines.

62 As laid out by Landes, *Wealth and poverty*, and by David E. Bloom and Jeffrey D. Sachs, ‘Geography, demography, and economic growth in Africa’, *Brookings Papers on Economic Activity*, 29, 1998, pp. 207–96.

many contesting causal claims in global economic history, and it is unfortunate if the measurement method by itself lends support to one line of interpretation. Factor endowments may indeed have a direct impact on the feasibility of aggregating the GDP measure – or, in the words of economists, geography is endogenous to the level estimates. This in turn precludes, or at the very least biases, the use of this measure in discussions regarding whether factor endowments are important exogenous causal factors to the standard of living.

Factor endowments do affect the national account estimates in ways that are not captured by correcting for differences in prices and consumption, and also because the distinction between observed and non-observed production varies across different localities.⁶³ It is widely recognized that factor endowments have had a large bearing on the choice of crops, agricultural techniques, and the formation of agrarian production systems and patterns of specialization, which in turn have influenced the ‘choice’ of institutions and state formation.⁶⁴ Sugihara argues that the development paths taken by the ‘East’ and the ‘West’ were radically different, and that typically the latter was capital-intensive while the former was labour-intensive.⁶⁵ Francesca Bray earlier reasoned along similar lines with a different emphasis.⁶⁶ In *Rice Economies* she argues that technological development in rice-growing economies was typically land- and labour-intensive, rather than land-extensive and capital-intensive. Rice-growing areas could hold larger population densities, and, while tractors were suitable capital inputs in wheat fields, they were not suitable for rice fields. Chandavarkar investigated patterns of industrialization in India before 1947, and argued that previous accounts had failed to recognize dynamic change in this sector because their vision had been blurred by a Eurocentric focus on large-scale, capital-intensive operations.⁶⁷

The problem is that the assumptions made in historical national accounting are not justified using the principles of reciprocal comparison. As already noted, when we step outside the period of official recording, these are estimates and not data *per se*. When the estimates are made, the data compilers increasingly have to rely on proxies and assumed relationships in whatever data are available in order to aggregate the accounts. A good example is provided by Prados de la Escosura and Alvarez-Nogal, who suggest some conjectural estimates for the economic growth of Spain between 1500 and 1850.⁶⁸ Previous

63 The International Comparison Project uses market prices, instead of farm gate prices, when creating the PPP terms, thereby overestimating the poverty of all those who consume commodities that they themselves produce.

64 The most influential and recent version of this argument was put forward by Kenneth L. Sokoloff and Stanley L. Engerman, ‘History lessons: institutions, factor endowments, and paths of development in the New World’, *Journal of Economic Perspectives*, 14, 3, 2000, pp. 217–32.

65 Kaoru Sugihara, ‘The East Asian path of economic development: a long-term perspective’, in Giovanni Arrighi, Takeshi Hamashita, and Mark Selden, eds., *The resurgence of East Asia: 500, 150 and 50 year perspectives*, Routledge: London, 2003, pp. 78–123.

66 Francesca Bray, *The rice economies: technology and development in Asian societies*, Berkeley and Los Angeles, CA: University of California Press, 1986; or eadem, ‘Patterns of evolution in rice-growing societies’, *Journal of Peasant Studies*, 11, 1984, pp. 3–33.

67 Rajnarayan Chandavarkar, ‘Industrialization in India before 1947: conventional approaches and alternative perspectives’, *Modern Asian Studies*, 19, 3, 1985, pp. 623–68.

68 Carlos Alvarez-Nogal and Leandro Prados de la Escosura, ‘The decline of Spain (1500–1850): conjectural estimates’, *European Review of Economic History*, 11, 3, 2007, pp. 319–66.

estimates had mainly relied on grain production data, for which tithe data are available, but data on manufacturing and services are not available. This was resolved by the authors by choosing to 'proxy output trends in industry and services through changes in urban population (adjusted to exclude those living on agriculture), so tendencies in total output and output per head can be established at regional and national levels'.⁶⁹ Such an approach can be deemed plausible or not, depending on whether manufactures and services were mainly urban-based throughout this period of Spanish history.

Referring to Prados de la Escosura and Alvarez-Nogal, Jean-Pascal Bassino and his co-authors adapt the same method when attempting to create measures of economic growth for Japan for the period 730–1872.⁷⁰ To their credit, the authors correctly note that their measures may neglect rural production, but this neglect may have more important implications than they assume. Their assumption that non-agricultural production was inversely related to rural population runs counter to the prevailing consensus in Japanese economic history, where relatively late urbanization has been explained by the prevailing strength of rural industries.⁷¹ It was not until the silk industry ultimately failed in the 1930s that this dominant characteristic of the Japanese economy changed and modern manufacturing began to go hand in hand with urbanization.⁷² The particular path of Japanese growth before the First World War is not picked up, however, by Bassano's use of urban population as a measure of non-agricultural production.

A similar issue of different adaptations to geography explains why Broadberry et al. find that English GDP per capita was so much higher than previously estimated. Maddison's estimate was that GDP per capita for Britain in the year 1500 was 714 international dollars, whereas the revised and preliminary estimate presented by Broadberry et al. was 1128 for the same year.⁷³ The basic reason for raising the GDP estimates, which puts England on a much higher income than those estimated for parts of Asia, is the particular character of agricultural production in England and Britain. Broadberry and his fellow authors argue that, while the calorific consumption was not higher, the consumption basket had higher value-added components. This leads them to conclude that, 'contrary to the claims of the California School, Western Europe was on a very different path of development from Asia long before the Great Divergence, characterized by high value added, capital intensive and non-human energy intensive production'.⁷⁴

69 *Ibid.*, p. 320.

70 Jean-Pascal Bassino, Stephen Broadberry, Kyoji Fukao, Bishnupriya Gupta, and Masanori Takashima, 'Japan and the great divergence, 730–1872', unpublished paper presented at the Third European Congress on World and Global History, London School of Economics and Political Science, April 2011.

71 Penelope Francks, *Rural economic development in Japan: from the nineteenth century to the Pacific War*, New York: Routledge, 2006.

72 Debin Ma, 'The modern silk road: the global raw-silk market, 1850–1930', *Journal of Economic History*, 56, 2, 1996, pp. 330–55.

73 Broadberry et al., 'British economic growth', 2009 version, unpublished paper presented at the World Economic History Congress, August 2009, Utrecht, table 24. (The 2011 version of the paper does not provide estimates for Britain in the year 1500.)

74 Broadberry et al., 'British economic growth', 2011 version, p. 35.

However, Kenneth Pomeranz did not fail to recognize that Asian and European agrarian systems were different. Indeed, this is the very starting point for his reciprocal comparison.⁷⁵ Broadberry et al. argue that the living standard of the English was higher, as measured by GDP, because they had more livestock, ate more meat, drank more ale, and milled more grain per capita. But, as Pomeranz has already shown, this comparison only shows how China differed from England, and fails to ask the crucial reciprocal question. Pomeranz compares in both directions and finds that, in parts of Asia, fewer cultivators had to sacrifice land for the production of livestock for transport because they had more efficient waterways, rice could be consumed without milling, and the population relied on tea instead of ale and bean curd instead of meats, which are arguably both healthier and more cost-efficient.⁷⁶ As Pomeranz reasons, any statement claiming the supremacy of living standards of one group over another that is based on differences in local diets is shaky. The example referred to above shows how this misunderstanding can be carried into the national accounts. At face value it is only a matter of adding up the numbers, but in fact this method may carry serious regional bias when the aggregates are used to measure living standards.

In response to Sugihara's formulation of a particular Asian path of development, Gareth Austin suggested that the African pre-colonial and colonial economies followed quite a different path.⁷⁷ In Sub-Saharan Africa, labour and capital were scarce,⁷⁸ whereas land was relatively abundant. This meant that the typical technique for agricultural production was land-extensive and labour-saving.⁷⁹ It follows that returns to labour in the non-modern sector continued to be high into the post-colonial period. Writing about the lack of data about development in Nigeria in 1966, Stolper concluded the following: 'The absence of a Malthusian problem makes it illegitimate to neglect the so-called subsistence sector and to assume that any increase in output by "modern" sectors is a net addition to total product.'⁸⁰ Jeffrey Herbst has argued that this explains the relative weakness of African states.⁸¹ The basic argument is that when land is abundant, states are not needed to protect private property rights and conversely are not able to collect taxes. This has direct implications for national accounting. If taxes are not collected in kind or in cash for agricultural production, calculating the aggregate production in these areas becomes guesswork.

One of the most basic engines of growth and modernization is structural change, as formalized in the Lewis model. In this model, income growth arises from an 'unlimited supply of labour' from a rural sector that has a marginal productivity of labour close to zero and is therefore employed in the modern sector at subsistence wages. In this section

75 Pomeranz, *Great divergence*, p. 32.

76 *Ibid.*, pp. 32–6.

77 Gareth Austin, 'Resources, techniques and strategies south of the Sahara: revising the factor endowments perspective on African economic development, 1500–2000', *Economic History Review*, 61, 3, 2008, pp. 587–624.

78 Although Austin makes the crucial distinction that labour scarcity was subject to seasonal variation.

79 With some important exceptions where land usage was intensive, such as in Rwanda and Ethiopia.

80 Wolfgang F. Stolper, *Planning without facts: lessons in resource allocation from Nigeria's development*, Cambridge, MA: Harvard University Press, 1966, p. 21.

81 Jeffrey Herbst, *States and power in Africa: comparative lessons in authority and control*, Princeton, NJ: Princeton University Press, 2000.

we have noted some caveats regarding this basic model of modernization in the cases of Japan and India, as well as the generalizations that have been made about ‘Asia’ and ‘Africa’. The main lessons learnt are that the marginal productivity of labour in the rural sector was higher than zero and that small-scale production was central to this dynamic. Data are more likely to be missing for these sectors, and the resulting reliance on proxies such as urbanization or imports of capital goods will bias level estimates. This ‘missing observation’ bias is not randomly distributed. It will make the ‘European path’ look more valuable than other successful adaptations to different factor endowments and climatic conditions.

Production versus income: measurement in decentralized or weak states

In addition to the difficulties in comparing production and income created by the issues already noted, further problems arise because societies have different levels of centralization, specialization, and market integration. The GDP estimates are most commonly reached using either the production or the expenditure approach. In the latter approach, the statistical office adds up recorded expenditure or income of households, firms, and governments; in the former, it aggregates the value added for each sector. There is a large difference in the resulting estimates, and this becomes particularly important when comparing decentralized or weak states. To take a specific example: when the income method was applied to national accounting in colonial Northern Rhodesia, the African peasant contribution to national income was only considered when it was marketed through official channels. The colonial state was poorly integrated into the local economy, and relied on a decentralized and indirect form of rule for a very limited collection of taxes.⁸² Only from 1949 onwards was the figure of £5 million included in the value of national income for Northern Rhodesia to account for ‘subsistence production’.⁸³ However, the amount was unchanged in the accounts between 1949 and 1953, despite population growth and inflation in this period, which led to the assumption that the value of total food production from African producers was decreasing quite rapidly during this time. In independent Zambia, estimates were done according to the production method, applying a per capita guesstimate from the Food and Agricultural Organization (FAO). The per capita contribution per farmer increased fivefold.⁸⁴

The colonial legacy of low central state legitimacy remains with us today,⁸⁵ and as a result there is a very poor empirical basis for estimating production in most colonial and post-colonial societies. This problem has been exacerbated by the growing importance of

82 Leigh A. Gardner, ‘Decentralisation and corruption in historical perspective: evidence from tax collection in British Colonial Africa’, *Economic History of Developing Regions*, 25, 2, 2010, pp. 213–36.

83 Federation of Rhodesia and Nyasaland, *Monthly Digest of Statistics*, Salisbury, 1955.

84 Republic of Zambia, *National Accounts 1964–1967*, Lusaka.

85 Mahmood Mamdani, *Citizen and subject: contemporary Africa and the legacy of late colonialism*, Princeton, NJ: Princeton University Press, 1996; Pierre Englebert, *State legitimacy and development in Africa*, Boulder, CO: Lynne Rienner, 2001.

the ‘informal sector’ in the decades following structural adjustment, and a reduced role of the state in the south since the 1980s.⁸⁶ For example, Janet MacGaffey noted that the income of the Democratic Republic of Congo (DRC) is grossly understated in the official statistics. In 1991 she suggested that the real economy might be three times larger than that recorded by the official statistics.⁸⁷ This observation was made some time ago, but the reach of the central state in the DRC is unlikely to have improved much since.⁸⁸ The DRC may be thought to be an extreme example, but it has been documented that the extent of underestimation was of comparable size in Tanzania;⁸⁹ and as late as last year, following a revision, the income estimate of Ghana was increased from about US\$600 per capita to about US\$1100 per capita.⁹⁰ Thus, the estimate given in Table 2 of 218 international dollars per capita in the DRC for the year 2000 should not be interpreted literally. The estimate is not directly comparable with contemporary estimates for countries that have strong central states today, and even less comparable with the historical national accounts wherein the authors make allowances for missing data by using proxies for the unrecorded economy.⁹¹

Both theoretical and empirical literature suggest that ‘state legitimacy’ has had a positive effect on economic development in different ways; however, one does not want to have a measure of development where the strength of the state, in terms of formal recording and taxation, appears on both sides of the equation, as when one is attempting to explain low levels of income with reference to relative weakness of states. The informal economy is perhaps testament to the weakness of the central state and, if GDP per capita is intended as an approximation of living standards, then a measure of recorded income – as opposed to actual disposable income – is inappropriate for comparison. This recommendation is in line with those made by the Commission on the Measurement of Economic Performance and Social Progress: ‘look at income and consumption rather than production’.⁹² It was argued that GDP mainly measures marketed production, but is often treated as a measure of economic wellbeing, and that ‘conflating the two can lead to misleading indications about how well off people are’.⁹³ This is easier said than done when it comes to historical national accounts. The problem increases in importance when we compare societies where the ability to record marketed production varies in a systematic manner – as it does when we compare decentralized states with centralized states.

86 Béatrice Hibou, *Privatizing the state*, New York: Columbia University Press, 2004.

87 Janet MacGaffey, *The real economy of Zaire: the contribution of smuggling and other unofficial activities to national wealth*, London: James Currey, 1991, p. 11.

88 Kristof Titeca and Tom de Herdt, ‘Real governance beyond the “failed state”: negotiating education in the Democratic Republic of the Congo’, *African Affairs*, 110, 439, 2011, pp. 213–231.

89 Morten Jerven, ‘Random growth’; idem, ‘Relativity of poverty’.

90 Ghana Statistical Service, *Rebasing of Ghana’s National Accounts to Reference Year 2006*, Accra, 2010.

91 Data provided in Table 1 above.

92 Joseph Stiglitz, Amartya Sen, and Jean Paul Fitoussi, ‘Report of the Commission on the Measurement of Economic Performance and Social Progress’, presented 14 September 2009, <http://media.ft.com/cms/f3b4c24a-a141-11de-a88d-00144feabdc0.pdf> (consulted 9 December 2011), p. 60.

93 *Ibid.*, p. 137.

The problems of monitoring and recording are not only related to state capacity itself. The challenge of recording is higher in some locations than it is in others. In an eloquent ‘prosecution’ of development economics, Polly Hill explains why official statistics in poorer economies in the south are in such a ‘calamitous state of affairs’.⁹⁴ In addition, other authors have noted that agricultural statistics from developing countries, and particularly from African countries, are poor.⁹⁵ The interest here is not in inaccuracies as such but in whether there are systematic biases relating to geography in the national accounting procedures. Hill reviewed the practices at the FAO, and noted that the instructions in the 1980 world census of agriculture contained no specific advice or treatment of the particular problems of tropical countries. They thus ignored the fact that different techniques such as crop mixtures and cross-cropping are not only very common but also more productive.⁹⁶ Hill argues that sampling methods were developed with modern, specialized, capital-intensive agriculture in the northern hemisphere in mind, and that these methods are not compatible with a non-temperate climate. Total agricultural productivity per crop gets complicated and the resulting figures are misleading. An example of this is provided by roots and tubers such as cassava or manioc often being planted with bananas and plantains. These crops, grown on tiny plots and in mixtures, are not harvested until they are needed, because the tubers only last for a couple of days. The FAO does provide statistics on these crops, but, as Hill argues, ‘no West African country can have the faintest idea as to how much is really produced’.⁹⁷

The lack of available data for full assessment of production and how it relates to inputs of land, labour, and capital has resulted in vigorous debates among practitioners of African economic history. In a special issue of *African Economic History* on pre-colonial agriculture and industry, John Thornton put forward the claim, supported by a few data observations, that ‘there was some basis to the earlier assertions that African agriculture, even without the plow, was more efficient than that of early modern Europe’.⁹⁸ This claim was rejected by other scholars in that same issue. In particular, Jan Hogendorn and Henry Gemery argued that this observation was misleading because the data quoted in the article related to yield per seed or yield per unit of land, while the constraining factor in the African context was labour.⁹⁹ In the same issue, E. Ann McDougall raised the issue of the integration of these pre-colonial economies, and wondered to what extent one-off observations on the quality or quantity of output in

94 Polly Hill, *Development economics on trial: the anthropological case for a prosecution*, Cambridge: Cambridge University Press, 1986.

95 See, for instance, Sarah Berry, ‘The food crisis and agrarian change in Africa: a review essay’, *African Studies Review*, 27, 2, 1984, pp. 59–112; Roy A. Carr-Hill, *Social conditions in Sub-Saharan Africa*, Basingstoke: Macmillan, 1990, esp. pp. 39–51.

96 Hill, *Development economics*, p. 34.

97 *Ibid.*

98 John Thornton, ‘Precolonial African industry and the Atlantic trade, 1500–1800’, *African Economic History*, 19, 1990–91, pp. 1–19.

99 Jan S. Hogendorn and Henry A. Gemery, ‘Assessing productivity in precolonial African agriculture and industry, 1500–1800’, *African Economic History*, 19, 1990–91, pp. 31–5, esp. p. 31.

agriculture and industry answer any questions when what we are interested in is the total production of the economy.¹⁰⁰

Herbert S. Frankel was one of the pioneers in providing national income estimates for colonial African economies. His objection to the idea of measuring incomes and comparing income across countries had a completely different basis. Consistent with ‘substantivism’,¹⁰¹ Frankel argued that some of the economic behaviour of Africans cannot be adequately explained by concepts drawn from market economics. According to him, societies have such different concepts of income and welfare and are governed by such specific rules and laws that international comparisons would be meaningless. The concept of income or wealth varies from culture to culture to such an extent that efforts to maximize it cannot be compared across them. Indeed, Frankel compared the maximization of income to that of maximizing a game of chess. A game of chess is governed by specific rules that set the aim of the game and, as such, the game cannot be maximized.¹⁰² Also arguing that the economics of scarcity is a systemic and specific concept, Marshall Sahlins suggested the thesis of ‘the original affluent society’.¹⁰³ This thesis is based on observations of hours worked and calories consumed in ‘primitive societies’, showing that the working week was shorter but caloric intake was higher than, for instance, the value used by Pritchett and Maddison for the lowest threshold of GDP per capita income in pre-modern societies. Sahlins thus used a non-cultural and non-market standard of comparison (hours worked and caloric intake) to argue that the economics of scarcity did not apply; thus production was not maximized, yet primitive societies were still affluent.¹⁰⁴

While one can disagree with the conclusions of Frankel and Sahlins, they raise questions about whether one can assume that societies aimed to maximize a similar production function through time and space. It is not certain that what is being ‘maximized’ is the same across societies; perhaps even more importantly, there is no unique strategy for how production is maximized. Generalizing from the materials discussed here, in pre-industrial agriculture societies in tropical Africa the scarce factors of production were labour and capital, and a rational actor would seek to maximize land-to-labour ratios. In rice-growing areas in Asia the equation had a quite different rational solution. Here land and capital was scarce, and thus the solution was the opposite: maximize labour-to-land ratios. In wheat-growing areas, and particularly those of the ‘New World’, the solution was to add more capital to land, thus saving on labour.

No longer do we think of markets and states as developing in a stage-like system such as that proposed by Karl Polanyi. Not only can the choices of production technique be explained as efficient outcomes and rational solutions in a given environment, but so can

100 E. Ann McDougall, ‘Production in precolonial Africa’, *African Economic History Review*, 19, 1990–91, pp. 37–43.

101 ‘Substantivism’ refers to the proposition, first put forward by Karl Polanyi, that both rational decision-making and the logic of economic scarcity are society-specific.

102 S. Herbert Frankel, ‘Psychic and accounting concepts of income and welfare’, *Oxford Economic Papers*, 4, 1, 1952, pp. 1–17.

103 Marshall David Sahlins, *Stone Age economics*, Chicago, IL: Aldine-Atherton, 1972.

104 For a discussion of the extent of ‘economic rationality’ in feudal systems, see Witold Kula, *An economic theory of the feudal system: towards a model of the Polish economy*, London: N.L.B., 1976, esp. chs. 6 and 7.

the creation of certain types of organization. The plough is an efficient solution in one place, but an inefficient solution in another. Typically, in temperate areas where land was relatively scarce and livestock available, the plough made sense, whereas in tropical areas where the soil fertility was low, land abundant, and cattle largely unavailable, it did not.¹⁰⁵ It may be argued that this in turn had bearings on the type of states that were formed. Jack Goody, for example, claimed that the lack of the plough explained the rarity of feudal societies in Africa.¹⁰⁶ This is where one enters the realms of grand causal narratives – a road laden with contesting arguments about origins, path dependencies, direction of change, and destination. The simple, yet crucial point made here is that the level estimates make empirical claims about the comparative standard of living achieved in different systems in the present and the past. Quite obviously, that judgement depends on the particular perspective not only of the researcher but also of people in the past, such as whether you were asking the serf or the lord. The information contained in the level estimates only captures the perspective of the person responsible for compiling the information in the present day. Therefore, at the furthest extreme, it amounts to telling history backwards. This occurs when our retrospectively created data (as in the income level estimates presented here) are shaped by presumed notions of the origin, paths, direction, and destination of the development that the data are supposed to describe.

Conclusion: writing global economic history with national accounts

As quoted in the introduction, Angus Maddison remarked that ‘quantification clarifies issues which qualitative analysis leave fuzzy’.¹⁰⁷ That statement seems intuitive, but this article suggests quite a different view. It is not at all clear that qualitative analysis is fuzzy while numbers are not. One of the desirable features of numbers is that their interpretation is straightforward – one will be larger or smaller than the other. This is only true, however, if we accept and trust the numbers as facts. This article has argued that there are specific spatial and temporal factors that condition the availability of historical sources. Thus, quantitatively important biases appear when a measure is applied and interpreted as a universal measure of development through time and space.

National income estimates are products rather than facts, and most obviously so outside the period of official national accounting. There is a considerable scholarly effort to improve and document the historical evidence that underlies the national income estimates for Europe and other regions before 1900. This scholarship tends to revise the estimates upwards. This article has made the point that such estimates may still be less clarifying than one could hope for because Maddison was correct in his assertion that they are ‘readily contestable’.

105 Hopkins, *An economic history of West Africa* and Austin, ‘Resources, techniques and strategies’.

106 Jack Goody, ‘Feudalism in Africa?’, *The Journal of African History*, 4, 1, 1963, pp. 1–18, and Jack Goody, ‘Economy and feudalism in Africa’, *The Economic History Review*, New Series, 22, 3, Dec., 1969, pp. 393–405.

107 Maddison, *Contours*, p. 1.

In order to create historical national accounts that are useful, one must overcome one basic bias. When states are less centralized, production is more rurally based, operations are less capital-intensive, and these activities are less likely to be recorded. Thus, when retrospective accounting is done, it is likely that these societies will appear to be poorer than they actually were at the time. The lack of data is not randomly distributed. Moreover, the data that are available may be weak and unreliable.¹⁰⁸ Assigning levels then becomes a matter of speculation and assumption. From the perspective of the producer the numbers are malleable, and thus for the user they may be seriously misleading.

It is therefore analytically useful to differentiate between situations in which national income estimates are ‘historical evidence’ in the sense that they were produced by official administrations, and those in which they are privately produced by independent scholars. The further the concept travels in time and space from its specific origin, the more sceptical we should be about its applicability to the society that is under investigation. Both need to be historicized and contextualized. This is particularly important for the use of national income estimates in global comparative history. The use of assumptions and proxies needs to be firmly grounded and justified with regards to the place that is being measured, rather than with respect to a global standard or path of development; the concept of reciprocal comparison needs to be fully employed. The lack of reliable data is systematically associated with states being less like European states, and it is therefore suggested that the use of national income estimates in global history results in an unlevel playing field.

*Morten Jerven is Assistant Professor at the School for International Studies at Simon Fraser University. He has published widely on national income accounting, African development statistics, and how these pertain to the debates on African economic performance both currently and over the long term. Recent publications include ‘African growth recurring: an economic history perspective on African growth episodes, 1690–2010’, *Economic History of Developing Regions*, 25, 2 2010, pp. 127–54, and ‘Users and producers of African income: measuring African progress’, *African Affairs*, 110, 439, 2011, pp. 169–190.*

108 Platt, *Mickey Mouse numbers*.